ELSEVIER

CrossMark

Journal of Clinical Epidemiology 67 (2014) 745–753

Journal of Clinical Epidemiology

ORIGINAL ARTICLES

Developing Core Outcome Measurement Sets for Clinical Trials: OMERACT Filter 2.0☆

Maarten Boers[a,*], John R. Kirwan[b], George Wells[c], Dorcas Beaton[d], Laure Gossec[e], Maria-Antonietta d'Agostino[f], Philip G. Conaghan[g], Clifton O. Bingham III[h], Peter Brooks[i], Robert Landewé[j], Lyn March[k,l], Lee S. Simon[m], Jasvinder A. Singh[n,o], Vibeke Strand[p], Peter Tugwell[q]

[a]*Department of Epidemiology and Biostatistics, VU University Medical Center, PK 6Z 165, PO Box 7057, 1007 MB Amsterdam, The Netherlands*
[b]*University of Bristol Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, United Kingdom*
[c]*Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada*
[d]*Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, Ontario, Canada*
[e]*Department of Rheumatology, Université Pierre et Marie Curie (UPMC), Paris 6, GRC-UMPC 08 (EEMOIS), and Assistance Publique Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, Paris, France*
[f]*Department of Rheumatology, Université Versailles-Saint Quentin en Yvelines, Ambroise Pare Hospital, Boulogne-Billancourt, France*
[g]*NIHR Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Leeds, United Kingdom*
[h]*Division of Rheumatology, Johns Hopkins University, Baltimore, MD 21224, USA*
[i]*School of Population and Global Health, University of Melbourne, Melbourne, Australia*
[j]*Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands*
[k]*Institute of Bone and Joint Research, Kolling Institute of Medical Research, University of Sydney, Sydney, Australia*
[l]*Department of Rheumatology, Royal North Shore Hospital, St. Leonards, New South Wales, Australia*
[m]*SDG LLC, Cambridge, MA 02138, USA*
[n]*Division of Rheumatology, University of Alabama at Birmingham, Birmingham, AL 35294, USA*
[o]*Veterans Affairs Medical Center, Birmingham, AL 35294, USA*
[p]*Division of Immunology/Rheumatology, Stanford University School of Medicine, Palo Alto, CA 94028, USA*
[q]*Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada*

Accepted 25 November 2013; Published online 28 February 2014

Abstract

Background: Lack of standardization of outcome measures limits the usefulness of clinical trial evidence to inform health care decisions. This can be addressed by agreeing on a minimum core set of outcome measures per health condition, containing measures relevant to patients and decision makers. Since 1992, the Outcome Measures in Rheumatology (OMERACT) consensus initiative has successfully developed core sets for many rheumatologic conditions, actively involving patients since 2002. Its expanding scope required an explicit formulation of its underlying conceptual framework and process.

Methods: Literature searches and iterative consensus process (surveys and group meetings) of stakeholders including patients, health professionals, and methodologists within and outside rheumatology.

Results: To comprehensively sample patient-centered and intervention-specific outcomes, a framework emerged that comprises three core "Areas," namely Death, Life Impact, and Pathophysiological Manifestations; and one strongly recommended Resource Use. Through literature review and consensus process, core set development for any specific health condition starts by identifying at least one core "Domain" within each of the Areas to formulate the "Core Domain Set." Next, at least one applicable measurement instrument for each core Domain is identified to formulate a "Core Outcome Measurement Set." Each instrument must prove to be truthful (valid),

This article presents the main arguments and supporting information on the development of OMERACT Filter 2.0. Further details are provided in a series of publications in the Proceedings of the OMERACT 11 conference, in press at the *Journal of Rheumatology*. These will be made freely available (see www.omeract.org).

Declaration of authorship and conflict of interest: All authors declare to meet the conditions for authorship. Each has made a substantial contribution to conception and design of the study, data acquisition, data analysis, and interpretation. M.B., J.R.K., and P.T. wrote the initial draft, all others revised the article for important intellectual content. All approved the final version of the submitted manuscript. All authors declare no competing interests.

* Corresponding author. Tel.: +31-20-444-4474; fax: +31-20-444-4475.

E-mail address: eb@vumc.nl (M. Boers).

0895-4356
http://dx.doi.org/10.1016/j.jclinepi.2013.11.013

discriminative, and feasible. In 2012, 96% of the voting participants ($n = 125$) at the OMERACT 11 consensus conference endorsed this model and process.

Conclusion: The OMERACT Filter 2.0 explicitly describes a comprehensive conceptual framework and a recommended process to develop core outcome measurement sets for rheumatology likely to be useful as a template in other areas of health care.

Keywords: Outcome assessment (health care); Clinical trial; Rheumatology; Research design; Biological markers; Reference Standards

1. Introduction

Randomized controlled trials (RCTs) are performed to compare benefits and harms between interventions. To inform health care decisions, outcome measures relevant to patients and decision makers must be included and reported [1].

Disagreement on the choice of outcome measures has resulted in inconsistent reporting, potential for reporting bias, and reduced quality of guidelines depending on such trial results [2,3]. Recent examples include chronic obstructive pulmonary disease [4], diabetes mellitus [5], and surgical oncology [6]. The Patient-Centered Outcomes Research Institute (PCORI) was established by the US Congress to improve the quality of information for health care decisions. It recommends ''measuring outcomes that people in the population of interest notice and care about'' [7]. Within rheumatology, the Outcome Measures in Rheumatology initiative (OMERACT) has successfully worked since 1992 to improve outcome measurement for many rheumatologic conditions [8], including rheumatoid arthritis [9], ankylosing spondylitis [10], and osteoarthritis [11]. It has done so by developing widely endorsed ''core outcome measurement sets,'' each a minimum set of outcome measures that must be reported in all RCTs in a given health condition. Reporting consistency is ensured when investigators always report the results of the core set measures, regardless of the primary purpose of the trial. The core sets by no means limit investigators in their choice of primary or other outcome measures of interest.

Core sets have also been developed in other health areas. Their development most frequently started with a literature review of trials to date for consideration by an expert panel of health care professionals [12]. Relevant domains for assessment were then selected, often through ranking, and sometimes including suggestions for instruments to measure these domains, to form the proposed core set [12].

Criticisms of such an approach include: first, that it does not include a systematic survey of stakeholders, especially patients; second, that it is not explicitly linked to a conceptual framework describing the concepts to be measured; and third, that instrument selection is not based on systematic evidence of their measurement properties. Thus, these core sets may omit important concepts (either overlooked or lost in the selection process) and recommend suboptimal instruments.

In contrast, the OMERACT consensus process is grounded in a framework first formulated by Fries et al. [13] and expanded by Kirwan [14], and a process of documenting the validity of selected instruments by applying the ''OMERACT Filter'' to each candidate instrument (Table 1) [15]. The Filter was published in 1998 and summarizes key instrument properties (validity, reliability and responsiveness, and usability) in three plain language words, namely Truth, Discrimination, and Feasibility. In addition, an expanding group of patients with a rheumatic condition have been selected and trained in measurement. They have participated in OMERACT conferences and working groups since 2002 to identify important domains and to ensure that these are appropriately addressed.

The OMERACT community has grown in membership and in number of conditions covered, making clarification and optimization of its framework and process necessary. This includes the distinction between potential domains (''what to measure'') and measurement instruments (''how to measure''); and the process to identify these and to reach consensus on which to include in a core set, procedures implicit in its earlier work but not overtly defined or described. OMERACT feels such work might also be of use in other areas of health care, for example in the recent ''Core Outcome Measures in Effectiveness Trials'' (COMET) initiative, which aims to convene core set developers across all disease areas [16].

This article describes the upgrade of the OMERACT framework and the process of development of core outcome measurement sets in rheumatology, collectively termed ''OMERACT Filter 2.0.'' We believe that this framework and process will be useful as a template for development of core sets in other subspecialties.

2. Methods

A systematic literature review was performed to describe the process by which available conceptual frameworks (or models) of health, disease, and disability arose, with a goal to inform and improve the implicit OMERACT framework [12–14]. From the review, it became apparent that other existing frameworks were not focused on measurement in studies of efficacy and effectiveness, and that their development process was insufficiently documented. As a consequence, the OMERACT executive decided to build a generic model of core set development around its own framework aimed at RCTs and longitudinal observational studies for endorsement by the OMERACT community.

Definitions of key concepts were collected from the literature and operationalized, or developed de novo.

What is new?

- The conceptual framework encompasses the complete content of what is measureable in a trial, including both patient-centered and intervention-specific information.
- The framework comprises three core 'Areas': Death, Life Impact and Pathophysiologic Manifestations; and one strongly recommended: Resource Use.
- Core set development for any specific health condition starts by identifying at least one core 'Domain' within each of the Areas to formulate the 'Core Domain Set'.
- Next, at least one applicable measurement instrument for each core Domain is identified to formulate a 'Core Outcome Measurement Set'.
- For applicability, each instrument must prove to be truthful (valid), discriminative, and feasible.

Within the Cochrane Collaboration, experts helped develop and gave feedback on subsequent drafts by way of a targeted survey supplemented with open comments. A total of 18 of these experts met at the second COMET conference in July 2011 for further discussions. Next, an extensive internet-based survey was conducted with three groups, namely COMET-2 participants, participants of the current and previous OMERACT conferences including patient experts, and subscribers to the Evidence-Based Health Listserv [17].

At the same time, a proposal for specification of the OMERACT process to develop core outcome measurement sets was circulated to all active working groups ($n = 20$) within OMERACT and adapted after feedback. Finally, at the OMERACT 11 conference in May 2012 in Pinehurst, NC, a presentation of the draft conceptual framework and the development process for core sets (collectively termed "OMERACT Filter 2.0") was discussed in five interactive sessions that combined plenary and small group discussions. Based on all feedback received, the final proposal was presented and voted on in the closing session. Attendees included clinician–researchers (58%), full researchers (11%), industry researchers (13%), patient experts (9%), and statisticians (3%). Of note, several of the researchers present at OMERACT and several members of the OMERACT executive have expertise in health economics.

Table 1. The original OMERACT Filter to determine applicability of a measurement instrument in a setting [15]

Truth
Is the measure truthful, does it measure what is intended? Is the result unbiased and relevant? The word "truth" captures issues of face, content, and construct validity (As gold standards are often not available, criterion validity is mostly not tested).
Discrimination
Does the measure discriminate between situations of interest? The situations can be states at one time (for classification or prognosis) or states at different times (to measure change). The word "discrimination" captures issues of reliability and sensitivity to change.
Feasibility
Can the measure be applied easily, given constraints of time, money, and interpretability? The word "feasibility" captures an essential element in the selection of measures, one that may be decisive in determining a measure's success.

Abbreviation: OMERACT, Outcome Measures in Rheumatology.

3. Results

After several iterations, the initial draft proposal was provided as prereading material for 2,293 participants approached for the internet survey (1,484 Evidence Based Health listserv participants; 131 COMET 2 participants; and 678 OMERACT previous and current conference delegates) [18]. A total of 262 survey responses were received (11%), with wide support for the draft model and suggestions for improvement and clarification. Many responses raised similar issues, and new suggestions were not found in the final set of returns, suggesting saturation of information. Detailed discussions at OMERACT contributed further clarifications [19–23]. Of 125 voting participants, 120 (96%) endorsed the final proposal, which we describe in the following sections [24].

The first part of Filter 2.0, the framework, is designed to ensure comprehensiveness (content validity) of the core set by specifying all key aspects, to be termed "Areas," of a health condition (Fig. 1): three Areas that describe the "Impact of Health Conditions," specifically Death, Life Impact, and Resource Use; and the fourth Area that describes Pathophysiological manifestations. We posit that these Areas encompass the complete content of what is measureable in a trial, including both patient-centered and intervention-specific information. The OMERACT 11 participants endorsed that all Areas except Resource Use should always be addressed in clinical trials. The Areas, further described in the following sections, can be likened to large "containers" for the concepts of interest, herein termed Domains and Subdomains.

The second part is the process of core set development. To decide what to measure, (Sub)Domains will need to be specified within each Area to create the Core Domain Set. Once this step is complete, the question on how to measure can be answered by deciding on the specific measurement instruments for each (Sub)Domain to create the Core Outcome Measurement Set. Thus, a core set specifies at least one Domain in each Area and at least one valid instrument in each Domain. This is described further in process

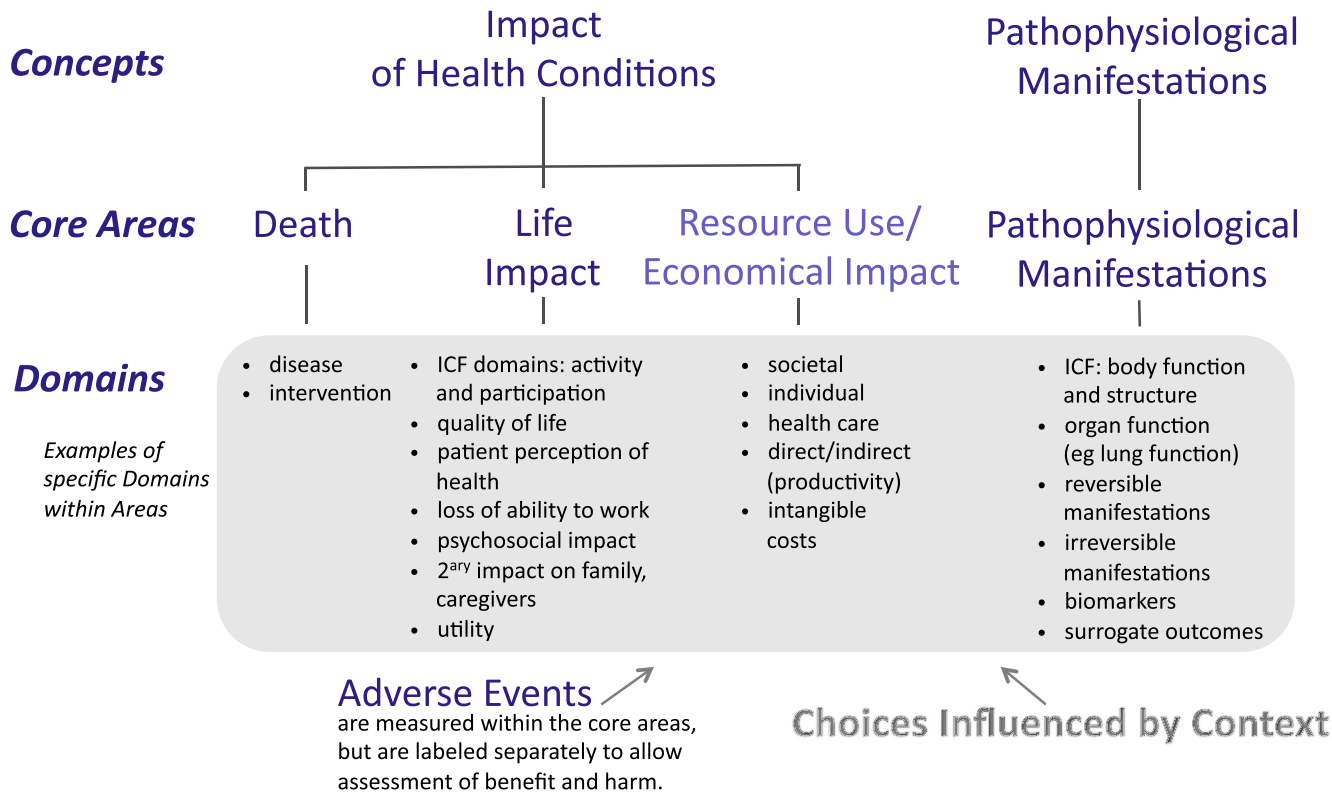

Fig. 1. Conceptual framework of Core Areas for outcome measurement in the setting of health intervention studies. Resource Use has a lighter shade to indicate it is currently strongly recommended, but not mandatory for inclusion. The choice of specific Domains within an Area depends on the context for which the core set is being developed in all areas, domains can be generic or made more specific, for example disease-specific, time-specific (eg, short or long-term), specific for patient preference, and so forth. ICF, International Classification of Functioning, Disability and Health.

suggestions in the following sections. The endorsed definitions of key concepts are listed in Table 2.

It is important to stress that a core set in no way limits the choice of primary and secondary outcomes in a particular trial. The core set describes the minimum set of Domains and Instruments to be measured in each of the Areas, which may already coincide with the primary or secondary outcomes of interest in that trial. If not, the investigator will have to include additional outcomes to comply with requirements of the core set.

4. The framework in further detail

4.1. Impact of health conditions

This includes all aspects of health or a health condition that are important to the patient and society. Impact is split into three Areas, namely Death, Life Impact, and Resource Use. In all Areas of Impact, there are various possibilities for further specification into "Domains" (and if useful into "Subdomains"); this will be determined by the condition and setting or scope for which the core set is being developed. Some suggestions for Domain specification are described in the following sections.

Under Death, possible specifications include generic and disease-specific, that is, all cause vs. disease-specific mortality; and intervention-specific (eg, death owing to surgery or transplantation). In conditions where death rarely occurs during a trial, this area could be covered in the core set by requiring a simple report of any deaths (or their lack), which is already a standard requirement in current guidelines.

Under Life Impact, OMERACT strongly suggests that core set developers consider both the domains of the International Classification of Functioning, Disability and Health (ICF) [25] and domains within the concept of health-related quality of life, for example, as elaborated by Wilson and Cleary's model [26]. The alignment of our Framework with ICF Domains and health-related quality of life is shown in Table 3. In trials primarily focused on understanding a mechanism of action or proof of concept, the core set will describe the minimum to be measured under Life Impact.

Resource Use describes the economic impact of health conditions both on society and on the individual. Both the presence of a health condition and its treatment incur resource use. In the development of health interventions, early consideration of resource use has become essential, as increasing health care costs present challenges for even the richest nations [27]. In low-income countries, the availability of an intervention and patient access to it may be determined by its associated resource use.

The OMERACT strongly recommends the inclusion of at least one Domain describing Resource Use, but it need

Table 2. Definitions of key concepts

Health
A state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity (WHO, 1948). Discussions are ongoing because this definition has many disadvantages, but it has not yet been replaced.
Health Condition
A situation of impaired health.
Health Intervention
An activity performed by, for, with or on behalf of a client(s) whose purpose is to improve individual or population health, to alter or diagnose the course of a health condition, or to improve functioning. (Draft WHO definition; 'by' added)
Core Area
An aspect of health or a health condition that needs to be measured to appropriately assess the effects of a health intervention. Core Areas are broad concepts consisting of a number of more specific concepts called Domains.
(Sub)Domain
Component of Core Area: a concept to be measured, a further specification of an aspect of health, categorized within a Core Area.
Outcome
Any identified result in a (Sub)Domain arising from exposure to a causal factor or a health intervention. (Adapted from John Last, *Dictionary of Epidemiology*. Toronto: Oxford Press 1995) Generic word that has been used with different definitions; has often been used interchangeably with ''Outcome Measure'' and ''Endpoint.''
Measurement Instrument
A tool to measure a quality or quantity of a variable, in this context a (Sub)Domain or a contextual factor. The tool can be a single question, a questionnaire, a score obtained through physical examination, a laboratory measurement, a score obtained through observation of an image, and so on.
Outcome Measurement Instrument
A measurement instrument chosen to assess Outcome. The result of measurement (recently termed 'specific metric' [33]) can be expressed as change, as end result, as cumulative result, or as or as ''time to event'' in a (Sub)Domain. Example: in pain measurement, the instrument could be a visual analog scale, and outcome could be an improvement on that scale (ie, change, an end-of study pain score, the achievement of an acceptable pain state (both, end result); or the area under the curve of pain scores during the study (cumulative result).
Core Domain Set
For studies of health interventions, the minimum set of Domains and Subdomains necessary to adequately cover all Core Areas, that is, adequately measure all relevant concepts of a specific health condition within a specified setting. Describes what to measure. Currently, the COMET initiative uses the term ''Core Outcome Set'' for this concept. OMERACT has decided not to adopt this term, as there is no consensus on its technical definition
Core Outcome Measurement Set
The minimum set of outcome measurement instruments that must be administered in each intervention study of a certain health condition within a specified setting to adequately cover a corresponding Core Domain Set. Describes how to measure. Within OMERACT, we have chosen to not to use the word ''endpoint,'' and use the word Outcome only in the context of the Core Outcome Measurement Set.
Setting (Scope)
The set of factors that describes the studies and circumstances to which the core outcome measurement set will apply. This is determined by the study questions and includes the health condition(s), target population, interventions, and so on.
Contextual Factor
Variable that is not an outcome of the study, but needs to be recognized (and measured) to understand the study results. This includes potential confounders and effect modifiers.

Abbreviation: OMERACT, Outcome Measures in Rheumatology.

not always be detailed. For instance, in a short trial of a new intervention, the measurement of resource use could be limited to listing the number of procedures and clinic visits necessary when this intervention is introduced in routine care. As more experience is gained, this Area may be changed from strongly recommended to Core and thus be addressed in all trials.

4.2. Pathophysiological manifestations of health conditions

In addition to Impact, OMERACT feels that in trials, measurement of pathophysiological manifestations is essential to assess whether or not the effect of the intervention specifically targets the pathophysiology of the health condition. Pathophysiology can include psychosocial manifestations. Example Domains include: organ function (eg, renal function), reversible manifestations (including modifiable risk factors and actual manifestations of ill health), and irreversible manifestations (including unmodifiable risk factors and damage). This Area also encompasses all biomarkers and surrogate outcomes. In trials primarily focused on Impact, the core set will describe the minimum to be measured under pathophysiological manifestations.

4.3. Additional concepts and considerations

4.3.1. Adverse events

Benefit and harm can be regarded as opposite directions on one ''impact scale,'' measurable in one of the core

Table 3. Comparison of OMERACT Framework to a model of Health-related Quality of Life [26] and the International Classification of Functioning, Disability and Health (ICF) [25]

OMERACT framework core areas/concepts	Health-Related quality of life model	ICF
Death		
Pathophysiological manifestations	Biological and physiological variables	Body function and structure
Life impact	Symptom status	Activity
	Functional status	
	General health perceptions	Participation
	Overall quality of life	
Contextual factors/scope	Characteristics of the individual	Contextual factors
	Characteristics of the environment	
	Nonmedical factors	
Health condition/scope		Health condition

Abbreviations: OMERACT, Outcome Measures in Rheumatology; ICF, International Classification of Functioning, Disability and Health.

Areas. However, current trials do not quantify harm (adverse effects) as carefully as benefit. Thus, OMERACT endorses that adverse events continue to be labeled or flagged to allow separate assessment of anticipated benefit and potential harm.

4.3.2. Setting and contextual factors

Core set developers need to specify the setting of the core set, and consider if any contextual factors need to be documented in every trial. Setting (or scope) includes the health condition, target population for the intervention, type of intervention, and so on. Contextual factors can be defined as those that are not the primary object of research but that may influence the results or the interpretation of the results. These include potential confounders and effect modifiers (most of which should be eliminated by randomization), as well as factors that define the generalizability of the study findings. One way of representing the interaction between contextual factors and other measured aspects of the impact of a health condition has been illustrated by the so-called "impact triad" [28].

5. Developing a Core Outcome Measurement Set: process suggestions

OMERACT suggests a stepwise approach to core set development (Figs. 2 and 3). The first step is defining the setting of the core set and deciding which (if any) contextual factors need to be measured alongside the outcome measures. Developers must also decide whether specific adverse events need to be monitored as part of the core set.

The next step is determining what to measure (Fig. 2), starting with a literature search to document all (sub)domains and instruments used to date. At the same time,

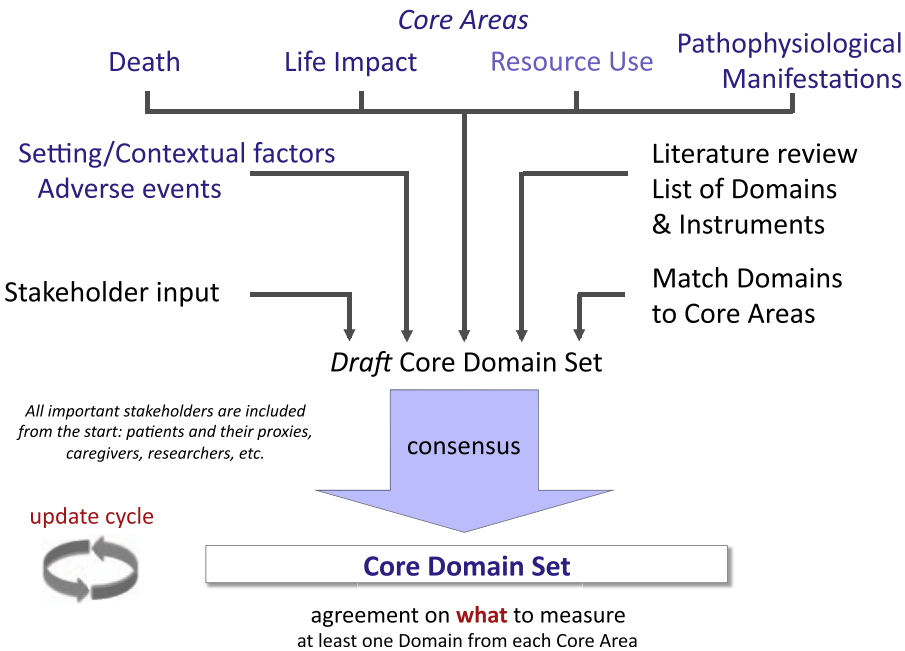

Fig. 2. Development of a Core Domain Set from the Core Areas of measurement. A Core Domain Set is defined as the minimum set of Domains and Subdomains necessary to adequately cover all Core Areas, that is, fully measure all relevant concepts of a specific health condition within a specified setting.

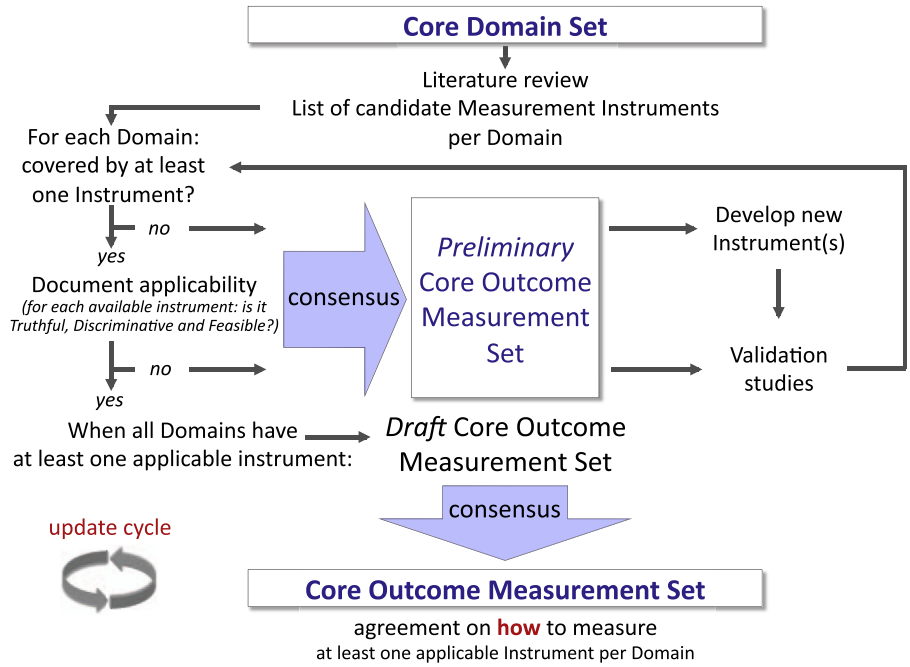

Fig. 3. Development of a Core Outcome Measurement Set from a Core Domain Set. The Core Outcome Measurement Set is defined as: the minimum set of outcome measurement instruments that must be administered in each intervention study of a certain health condition within a specified setting to adequately cover a corresponding Core Domain Set. As depicted, the development process allows core set developers to declare a Preliminary Core Outcome Measurement Set when not all Domains are covered by at least one applicable measurement instrument.

developers initiate stakeholder consultation to determine what each stakeholder group deems essential to measure. During this process, developers refer to the framework and match the input to specific Domains and Subdomains in each Core Area relevant to the chosen setting. To ensure face and content validity, explicit input from all stakeholders, including patients, is essential to identify relevant (sub)domains and to expose gaps in what has been measured to date [29]. The end result of this process is a draft subjected to a consensus procedure with all stakeholders, resulting in a Core Domain Set. Core Domains in the Areas of Death, Life Impact, and Resource Use can be equated with the "critical outcomes" of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system [30]. To do more justice to the value of all the domains and instruments reviewed, these can be viewed as occupying a series of concentric spheres: the core domains in the middle, surrounded by domains of decreasing importance [11].

The final step is deciding how to measure the selected Core Domains (Fig. 3). The literature review provides a list of available measurement instruments; where no instruments are available in a (sub)domain, these need to be developed. Each instrument is then studied to document its applicability in the chosen setting (Table 1) [15]. Drawing on key developments that have occurred after the formulation of the first Filter such as the work by the COSMIN group [31,32], and the GRADE recommendations [30], OMERACT is developing specific guidelines on the procedures to document applicability of instruments.

When all Core Domains can be measured by at least one applicable instrument, the end result is a draft that again is subjected to a consensus procedure with all stakeholders, resulting in the Core Outcome Instrument Set. Where core contextual factors and adverse events have been specified, instruments to measure these must also pass the OMERACT Filter. Formulation of a "preliminary" Core Outcome Measurement Set can be useful in situations where one or more Domains were incompletely covered by applicable measurement instruments.

6. Discussion

OMERACT has been active in the field of outcome measurement in rheumatology since 1992. It has now upgraded and clarified its working process for developing core outcome sets in two ways. First, it has formulated a novel conceptual framework of measurement of health conditions in the setting of health interventions. The Core Areas do not describe novel concepts, but their explicit juxtaposition in the framework is novel; in this way they ensure content validity across patient-important and intervention-specific information. In its overarching concept of "Impact," the framework confirms that assessment of patient-centered outcomes in trials is essential to inform health care decisions, as suggested by the PCORI [1]. Second, OMERACT has made the process of subsequent outcome development explicit by agreeing on the definitions of key concepts, including that of a Core Domain Set that logically must be developed before a Core Outcome Measurement Set.

The OMERACT Filter 2.0 identifies outcome domains and measurement instruments that will be consistent with the reporting framework recently developed for the US Clinical Trial Registry [33]. The registry additionally defines the term ''specific metric'' for the expression of the measurement result in a study (eg, as change score or as end score), and ''method of aggregation'' for the way the measurement result is summarized for the patients in each treatment group (eg, as mean and standard deviation).

During the development of this framework and process, now collectively termed ''OMERACT Filter 2.0,'' we sought to be maximally inclusive by engaging as many stakeholders as possible. OMERACT has recognized the patient perspective since its initiation [34] and included patients educated in measurement methodology in its working process since 2002 [35].

Energized by the enthusiastic reception and feedback received from participants of the COMET conference, we strove to make the framework as generic as possible so that it could be applied in other disease areas. Consequently, it incorporates considerable input from professionals working in other medical specialties. Our survey intended to reach out to a wide audience but many of the people approached (especially from the Evidence Based Health Listserv [17]) chose not to respond. This is a common experience with internet-based surveys, in this case probably caused both by lack of interest and the burden of the task. At this stage of development, we suggest that the quality and saturation of information is more important than a high response rate. Nevertheless, the framework is grounded in the setting of chronic diseases, and more specifically, rheumatology; the low overall response rate to the survey precludes conclusions on the acceptability of the framework in the wider scientific community. At present, the framework has face validity through its development process and the high degree of consensus obtained. This will need to be complemented by evidence of applicability in use. As such, the term ''Filter 2.0'' indicates the possibility of iterative modification over time as experience with its application expands and gaps and/or redundancies are identified. The next step is the development of an OMERACT Handbook with updated and explicit working methods to select and assess the quality of measurement instruments that will populate a core outcome measurement set.

In conclusion, OMERACT has established a strongly endorsed, comprehensive framework and process for developing core outcome measurement sets, which has been fruitful within rheumatology. We hope that OMERACT Filter 2.0 will serve a more generic function, and be applied in the development of core outcome measurement sets for many health conditions.

Acknowledgments

The authors thank Caroline Terwee, Leanne Idzerda, Reuben Escorpizo, Annelies Boonen, Susan Magasi, Ian Sinha, Paula Williamson, Jane Blazeby, and Gerold Stucki who gave valuable advice on this manuscript. They also thank all survey participants.

References

[1] Gabriel SE, Normand SL. Getting the methods right—the foundation of patient-centered outcomes research. N Engl J Med 2012;367:787−90.

[2] Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. Updated March 2011. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Intervention Version 510. The Cochrane Collaboration; 2011. Available at http://www.cochrane-handbook.org. Accessed December 30, 2013.

[3] Williamson PR, Altman D, Blazeby JM, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. J Health Serv Res Policy 2012;17:1−2.

[4] Cazzola M, MacNee W, Martinez FJ, Rabe KF, Franciosi LG, Barnes PJ, et al. Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur Respir J 2008;31:416−69.

[5] Gandhi GY, Murad MH, Fujiyoshi A, Mullan RJ, Flynn DN, Elamin MB, et al. Patient-important outcomes in registered diabetes trials. JAMA 2008;299:2543−9.

[6] Blencowe NS, McNair AG, Davis CR, Brookes ST, Blazeby JM. Standards of outcome reporting in surgical oncology: a case study in esophageal cancer. Ann Surg Oncol 2012;19:4012−8.

[7] Patient-Centered Outcomes Research Institute. Preliminary draft methodology report; 2012. Available at http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf. Accessed December 30, 2013.

[8] OMERACT website. Available at http://www.omeract.org. Accessed November 11, 2012.

[9] Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World health organization and international league of associations for rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol 1994;21(suppl 41):86−9.

[10] van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. J Rheumatol 1997;24:2225−9.

[11] Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand OA. Consensus development at OMERACT III. J Rheumatol 1997;24:799−802.

[12] Idzerda L, Rader T, Tugwell P, Boers M. Can we decide which outcomes should be measured in every clinical trial? Systematic review of concepts and methodologies used to establish core sets for outcome measurement. J Rheumatol 2014; In press.

[13] Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982;9:789−93.

[14] Kirwan JR. A theoretical framework for process, outcome and prognosis in rheumatoid arthritis. J Rheumatol 1992;19:333−6.

[15] Boers M, Brooks P, Strand V, Tugwell P. The OMERACT Filter for outcome measures in rheumatology. J Rheumatol 1998;25:198−9.

[16] Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials 2012;13:132.

[17] Evidence based health. Available at https://http://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=evidence-based-health. Accessed October 19, 2012.

[18] Boers M, Idzerda L, Kirwan JR, Beaton D, Escorpizo R, Boonen A, et al. Towards a generalized framework of core measurement areas in clinical trials: a position paper for OMERACT 11. J Rheumatol 2014;41:978−85.

[19] Kirwan JR, Boers M, Hewlett S, Beaton D, Bingham CO, Choy E, et al. Updating the OMERACT Filter: core Areas as a basis for defining core outcome sets. J Rheumatol 2014;41:994−9.
[20] Tugwell P, Boers M, D'Agostino M, Beaton D, Boonen A, Bingham CO, et al. Updating the OMERACT Filter: Implications of Filter 2.0 to select outcome instruments through assessment of content, face and construct validity. J Rheumatol 2014;41:1000−4.
[21] Wells GA, Tugwell P, Boers M, Kirwan JR, Beaton D, Landewé R, et al. Updating the OMERACT Filter: Discrimination and Feasibility. J Rheumatol 2014;41:1005−10.
[22] Kirwan JR, Bartlett SJ, Beaton D, Boers M, Bosworth A, Brooks PM, et al. Updating the OMERACT Filter: Implications for patient reported outcomes. J Rheumatol 2014;41:1011−5.
[23] D'Agostino MA, Boers M, Dougados M, Van Der Heijde D, Iagnocco A, Landewe RBM, et al. Updating the OMERACT Filter: Imaging and soluble biomarkers. J Rheumatol 2014;41:1016−24.
[24] Boers M, Kirwan JR, Gossec L, Conaghan PG, D'Agostino M, Bingham CO, 3rd, et al. How to choose core outcome sets for clinical trials: OMERACT 11 approves Filter 2.0. J Rheumatol 2014;41:1025−30.
[25] International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization; 2001:Available at. http://www.who.int/classifications/icf/en. Accessed October 22, 2012.
[26] Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA 1995;273:59−65.
[27] Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health 2005;8:521−33.
[28] Sanderson TC, Hewlett SE, Flurey C, Dures E, Richards P, Kirwan JR. The impact triad (severity, importance, self-management) as a method of enhancing measurement of personal life impact of rheumatic diseases. J Rheumatol 2011;38:191−4.
[29] Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174−7.
[30] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924−6.
[31] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737−45.
[32] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539−49.
[33] Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database—update and key issues. N Engl J Med 2011;364:852−60.
[34] Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol 1993;20:557−60.
[35] Kirwan J, Heiberg T, Hewlett S, Hughes R, Kvien T, Ahlmen M, et al. Outcomes from the patient perspective Workshop at OMERACT 6. J Rheumatol 2003;30:868−72.